A chapter from PLANETARY SATELLITES edited by J A Burns, Univ of Arizona Press, Tucson Az, 1977

THE CRITICAL VELOCITY PHENOMENON AND THE ORIGIN OF THE REGULAR SATELLITES

Bibhas R. De, Hannes Alfvén, and Gustaf Arrhenius
University of California, San Diego

The three well-developed satellite systems of Jupiter, Saturn, and Uranus may hold the key to the origin of the solar system. The spatial groupings in these satellite systems lead us to postulate a physical phenomenon that may have been operative during the formation era of the satellites. This is the now well-known and experimentally verifiable critical velocity phenomenon. A similar spatial grouping is present also in the planetary system, possibly indicating a process of formation of the planets similar to that of the satellites. It is suggested, however, that in trying to explore the genesis of the solar system, the primary emphasis should be laid on the satellite systems rather than the planetary system–since this approach puts important constraints on any theoretical model. A coherent series of processes that may have led to the formation of satellites (and, as a corollary, the planets) is outlined. The main thrust of the discussion is the crucial role of plasma phenomena in the formative processes of the solar system.

We shall discuss a physical concept leading to a theory of origin of the regular satellites, that is, the planetary satellites that are likely to have formed *in situ*. Thus, we exclude from our discussion the Moon (Lyttleton, 1967; Gerstenkorn, 1955, 1968, 1969; Singer, 1968, 1970a, 1972; Alfvén and Arrhenius, 1972; cf. Wood, Chapt. 27 herein), Triton (McCord, 1966), Phoebe, and the Jovian satellites VI through XIII (Bailey, 1971, 1972) in view of the likelihood that these satellites may be captured bodies. Capture processes are described in Chapter 7 by Burns. The regular satellites of Mars will be excluded because of the rudimentary character of this system (cf. Burns, 1972; Pollack, Chapt. 14 herein). We shall also exclude the regular satellite Nereid of Neptune because this provides us with only one member of a satellite system and because of the possibility that the Neptunian satellite system may have been modified due to the capture of Triton (McCord, 1966).

In the course of our discussion we shall recognize that because of the striking similarity between the satellite systems and the planetary system, it is both logically proper and aesthetically appealing to try to examine the premise that these two systems have an identical genesis. Hence we shall keep the planetary system in the back of our minds as we try to develop a theory for the satellite systems.

Note: For a more detailed development of the theory outlined here, the reader is referred to the monograph by Alfvén and Arrhenius (1976).

BAND STRUCTURES

Band Structure in the Solar System

We start by defining a quantity

$$\Gamma = \frac{M_c}{R_o} \ , \tag{1}$$

where M_c is the mass of a central body (a planet or the Sun) and R_o is the orbital radius of a secondary body (a satellite or a planet). Thus G Γ is the specific gravitational potential energy of the orbiting secondary body, where G is the gravitational constant. In Figure 24.1 the quantity log Γ has been represented along the radius of a circle. In this diagram the radial distance represents the inward trajectory of an object falling freely towards the central body. In the left sector of this diagram we have drawn an arc for each of the satellites of the Jovian, the Saturnian and the Uranian systems. We notice that the satellites tend to fall in bands—one band (or possibly two) shortward of log Γ = 19 and one band between log Γ = 19 and 20. There is only one satellite (Amalthea) with log Γ greater than 20 and this may represent the single member of a third band. The planets Saturn and Uranus could not have any satellites in this band since the surfaces of these planets (as marked by broken lines in the diagram, Γ being now defined as the planetary mass divided by the planetary radius) lie outside this band.

In the right sector of the diagram we have included the planetary system in order to explore the similarities between this system and the satellite systems. Here again the planets fall in two bands—one containing the terrestrial planets and the other containing the Jovian planets. The terrestrial planets complement and confirm the suspected Amalthea band in the left sector. The Jovian band overlaps the two outer satellite bands. We have included the Moon and Triton along with the planets in view of the likelihood that these may be former planets.

Having observed this band structure, we must now ask if the grouping of secondary bodies in the satellite systems as well as in the planetary system into discrete ranges of Γ is physically significant. To answer this question, we shall look for a possible physical explanation of this grouping.

Band Structure for Cosmically Abundant Elements

Suppose that a cloud of gas containing atoms of mass m is falling freely towards a central body. Let v and R be the velocity of infall of the cloud and its distance from the central body at any instant of time. If we neglect the thermal velocity of the atoms compared to their velocity of infall, then the kinetic energy of an atom at any instant is

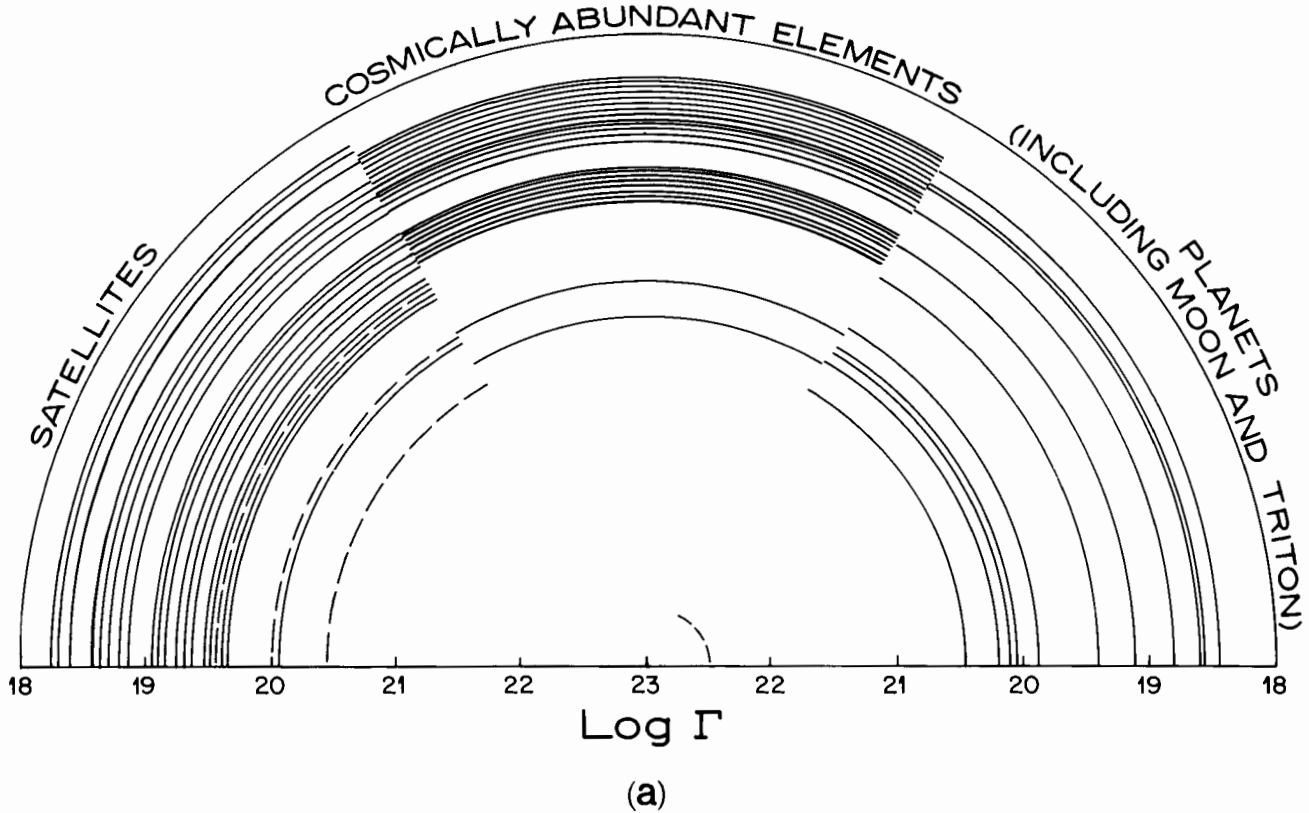

Fig. 24.1. (a) The band structure in the satellite systems, the planetary system and the cosmically abundant elements.

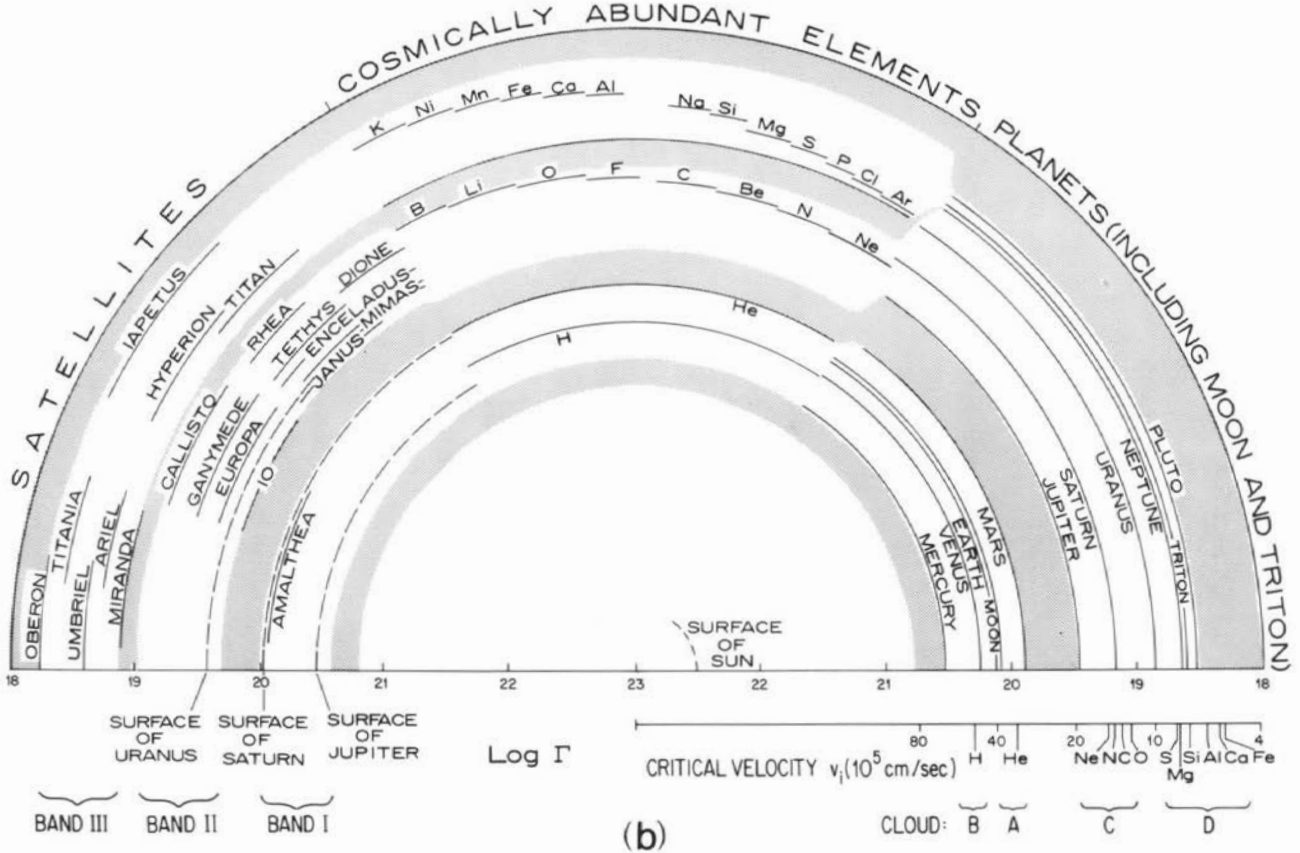

Fig. 24.1. (b) Key for the preceding diagram. This figure also includes a scale for the critical velocity for various elements, and the A, B, C and D clouds are indicated. The shading highlights the band structure.

$$\frac{1}{2}mv^2 = \frac{GM_c m}{R} . \tag{2}$$

Suppose that when the atom reaches a distance $R = R_i$ ("the ionization distance") and a corresponding velocity $v = v_i$, its kinetic energy becomes equal to its ionization energy eV_i (e = electronic charge, V_i = ionization potential). Hence we can write

$$eV_i = \frac{1}{2}mv_i^2 = \frac{GM_c m}{R_i} . \tag{3}$$

We may now define a parameter Γ_i for the infalling cloud (analogous to Γ for the satellites) by

$$\Gamma_i = \frac{M_c}{R_i} = \frac{eV_i}{Gm}$$

The parameter Γ_i for a cloud of infalling atoms is defined in terms of fundamental physical constants and is therefore itself a physical constant. In the middle sector of Figure 24.1 we have plotted the quantity $\log \Gamma_i$ for the first 23 cosmically abundant elements. We find that *these elements divide themselves into three distinct bands: Band I containing the elements in the first row of the periodic table, Band II containing the elements in the second row and Band III containing the elements in the third and the fourth rows.* Since this is a trend in a physical constant for the elements, it represents a fundamental property of matter.

Correspondence Between the Band Structures

On comparing the band structure for the elements in the middle sector of Figure 24.1 with the band structures in the left and the right sectors, we find that *the element bands coincide, substantially and without overlap, with the three satellite bands.*

In the planetary system, the terrestrial band coincides with Band I; the Jovian band coincides in part with Band II but partially overlaps Band III. The possibility arises therefore that the Jovian band has a correspondence to Band II but that it has shifted outward due to some reason that we should explore (for an explanation of this shift, see Alfvén and Arrhenius, 1974, 1975). The planet Pluto distinguishes itself from the Jovian sequence by having a higher density and orbital eccentricity, and hence it is possible that Pluto (and perhaps also Triton before its capture) has a correspondence to Band III. Neptune may have formed in a region of overlap between Bands II and III. In the following sections we shall use the band labels I, II and III to refer to the elements as well as to the satellites and the planets.

The correspondence between the band structures for the satellites and the

planets on one hand and the elements on the other may be stated in the following way: Γ_i for the elements falls into ranges which are identical to Γ for the satellites and the planets. Hence we may conclude that from an observational point of view there exists an identity between the distributions of the parameters Γ_i and Γ: thus we may put $\Gamma_i \equiv \Gamma$. Since $\Gamma_i = M_c/R_i$ and $\Gamma = M_c/R_o$, this identity translates into the following statement: The average orbital distance R_o of the secondary bodies in a band is identical to the ionization distance R_i of a cloud containing one or more of the elements in that band.

It is this identity whose explanation we shall seek.

THE CRITICAL VELOCITY PHENOMENON

Concept of Critical Velocity

How was the matter that now forms the regular satellites derived? This material could not possibly have been ejected from the planets (see Alfvén and Arrhenius, 1974). There remain two other possibilities: (1) the planets captured heliocentric grains which later accumulated to form the satellites, somewhat as suggested in Chapter 26 by Safronov and Ruskol, or (2) the grains condensed in the neighborhood of the planets from gases accumulated by the gravitation of the planets, as mentioned in Chapter 23 by Cameron. The orbital angular momentum of the grains could then derive from the planetary spin by a mechanism described later. This would require the eventual satellites to orbit in the prograde sense in the equatorial plane of their respective planets, regardless of how this plane is oriented with respect to the ecliptic. (It is possible that the gravitationally accumulating gas already contained interstellar grains. Such grains would behave as a part of the gas, resulting in a colloidal plasma when the gas becomes ionized on reaching its critical velocity. Angular momentum would be supplied to this colloidal plasma hydromagnetically, so that these grains would eventually be found as accumulated larger bodies orbiting in the equatorial plane of the planet.) We shall pursue the second alternative here. We shall further make the assumption that, during the era of formation of the satellites, the planets possessed dipolar magnetic fields—the axes of the dipole being roughly aligned with the planetary spin axes. In the case of the formation of the planetary system, the same assumption can be made with regard to the Sun.

Returning now to the correspondence between the ionization distance R_i of a cloud and the orbital distance R_o of a group of secondary bodies, we note that such a correspondence would be expected if the infalling cloud is stopped at the distance R_i, forming a group of satellites at this distance. Such a stopping of the cloud can occur at the distance R_i if the gas becomes ionized and the magnetic field of the planet prevents further infall. Since the kinetic energy of the gas atoms equals their ionization energy at this location, the gas can, in principle, become ionized if there exists a mechanism which can convert the kinetic energy of the atoms into their ionization energy.

If such an ionization mechanism exists, then the velocity v_i, defined by eqn. (3), i.e., by

$$v_i = \left(\frac{2eV_i}{m}\right)^{1/2} , \tag{5}$$

plays the role of a *critical velocity* such that when the velocity of a gas cloud reaches the value v_i appropriate to the gas, it becomes ionized (presumably by interacting with the background plasma and the magnetic field through which the cloud falls).

In this way, analysis of the orbital properties of the satellites and the planets, and its implications led to the concept of the critical velocity phenomenon (Alfvén, 1942–45, 1954), but an experimental demonstration of the phenomenon had not yet come.

The gas cloud we envisage as a precursor for a satellite group is a local cloud (''the source cloud'') at a large distance (larger than the planet-to-outermost satellite distance) from the planet. This cloud is assumed to be initially at rest relative to the planet. When the cloud has reached the ionization distance R_i, the thermal velocity of the atoms is assumed to be much smaller than v_i, so that the atoms will be falling radially inward. (For a hydrogen cloud, the requirement is that the temperature of the cloud be much less than 10^5°K at this stage.)

There is assumed to be a thin ''background'' plasma medium surrounding the planet—the plasma being suspended in the magnetic field lines. The neutral gas atoms fall through this plasma. The density of the plasma is assumed to be so low that the mean free path of the atoms is much larger than the radial extent of the plasma medium. Thus there are very few collisions between the plasma particles and the atoms of the infalling gas. In spite of this, the infalling gas is supposed to be rapidly ionized when the relative velocity between this gas and the magnetized plasma exceeds the critical velocity. The original existence of a thin plasma around a planet—in fact the existence of a thin plasma anywhere in space—can be taken for granted.

Under these conditions, clouds of neutral gas falling toward a planet would be stopped in bands appropriate to the chemical composition of the cloud (or in bands appropriate to the dominant or ''controlling'' element in the clouds) and form condensates which eventually accrete to form the observed groups of satellites. For the formation of the planets, the same mechanism would be envisaged to take place around the Sun.

Applicability of the Critical Velocity Concept to the Satellite Formation Problem

When the preceding analysis was first made, it had three difficulties: (a) The ionization mechanism described above was not known to exist. (b) There was no theoretical basis for the hypothesis that masses of gas falling in toward the central

bodies would have differing chemical compositions. (c) The chemical composition of the secondary bodies found in a band do not represent the chemical composition of the corresponding element band. Since practically nothing is known about the chemical composition of the satellites, let us turn to the planetary system for this discussion. For example, the terrestrial planets fall in a band which corresponds to the Γ values of H and He, but they contain very little of these elements. The band of the Jovian planets corresponds to C, N and O, but they were believed to consist mainly of H and He.

However, the above situation has changed drastically since the 1940s when it was first propounded because of both theoretical studies and empirical findings. Although we are still far from a final theory, it is fair to state that difficulty (a) has been eliminated by the discovery of the critical velocity phenomenon which is discussed in the next section. Difficulty (b) can be rationalized when we recognize that the separation of elements by plasma processes is a common phenomenon in space; such a separation process will be presented later. The apparent difficulty (c) is resolved by considerations given by Alfvén and Arrhenius (1976).

Experimental Verification of the Critical Velocity Phenomenon

In view of the fact that the theoretical treatment of plasma processes in space remains speculative unless supported by experiments, it was realized that further consideration of the applicability of the critical velocity concept to the problem of formation of satellites and planets depended on experimentally demonstrating the validity of this concept. As soon as the advance of thermonuclear technology made it possible, experiments were designed to investigate the interaction between a magnetized plasma and a non-ionized gas in motion relative to the plasma. These experiments have established beyond doubt the existence of the critical velocity phenomenon. A survey of these experiments has recently been given by Danielsson (1973), while a review of the various theories of the critical velocity phenomenon has been given by Sherman (1973).

The experiment most closely related to the cosmic situation was carried out by Danielsson (1969, 1970). The experimental arrangement is shown in Figure 24.2. A hydrogen plasma is generated and accelerated in an electrodeless plasma gun (a conical theta pinch). This plasma flows into a drift tube along a magnetic field. The direction of the magnetic field changes gradually from axial to transverse along the path of the plasma. A polarization electric field is developed and a plasma with a density of about 10^{11}–10^{12} cm^{-3} proceeds drifting across the magnetic field with a velocity up to 3×10^7 cm sec^{-1}. In the region of the transverse magnetic field the plasma penetrates into a small cloud of gas, which is released from an electromagnetic valve. This gas cloud has an axial depth of 5 cm and a density of 10^{14} cm^{-3} at the time of arrival of the plasma. The remainder of the system is under high vacuum. Under these conditions the mean free path

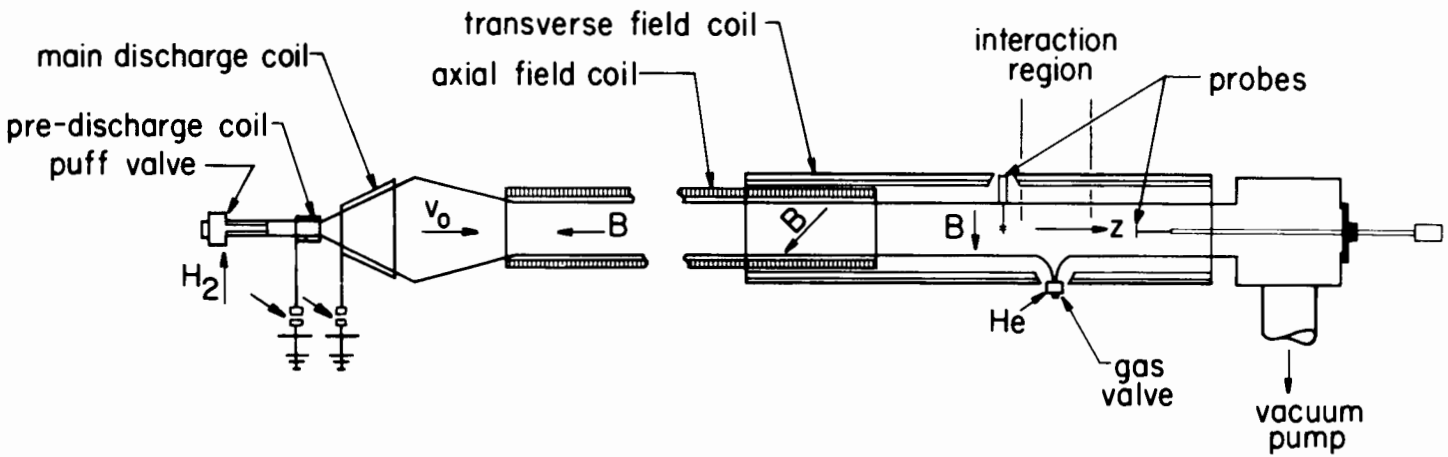

Fig. 24.2. The experimental arrangement of Danielsson (1969, 1970) for measuring critical velocity: The left part is a plasma gun, emitting a magnetized plasma with a velocity v_0. In a long drift tube, the longitudinal magnetization is changed to transverse magnetization. A thin cloud of gas is injected through the gas valve. If v_0 is below the critical velocity, the plasma beam passes through the gas cloud with very little interaction since the mean free path in the gas cloud is large. If v_0 is above the critical velocity, there is a strong interaction, decelerating the plasma beam and at the same time partially ionizing the gas cloud.

for direct binary collisions is much larger than 5 cm so that the interaction, according to common terminology, is collisionless.

The result of the experiment is that as long as the plasma velocity v_0 is below a certain value v_c, the plasma beam passes through the gas cloud with very little interaction, since the mean free path is large. If v_0 exceeds v_c, there is a strong interaction, bringing v_0 down to near v_c. At the same time the gas cloud becomes partially ionized. In general, the interaction leads to (a) local heating of the electrons, (b) ionization of the neutral gas and (c) deceleration of the plasma stream.

This and a number of other experiments have demonstrated the existence of a critical velocity v_c. Although under certain conditions there are deviations up to perhaps 50%, the general result is that v_c is the same as the velocity v_i given by eqn. (5).

THE FORMATION PROCESS

Chemical Composition of Infalling Clouds

Let us assume that the source cloud which contains all elements (having for instance abundances appropriate to some average "galactic" composition) is partially ionized to such an extent that all elements with ionization potentials higher than a certain value V_I are neutral, but all elements having ionization potentials lower than V_I are ionized. The Larmor radii of the electrons and ions are assumed to be much smaller, and the mean free paths much larger, than the dimensions of the source cloud. The neutral atoms will then begin to fall toward the central body.

Let V_I decrease slowly with time (e.g., through a magnetically induced cooling of the plasma as discussed by De, 1973b, in the case of solar prominences). When it has fallen below the ionization potential of helium, helium will recombine and a gas cloud dominated by helium will fall towards the gravitating central body. Helium reaches its critical velocity of 34.4×10^5 cm sec^{-1} at a Γ value of 0.9×10^{20}, that is, in the outer region of Band I in Figure 24.1. At this location the gas will become ionized and form a plasma cloud which will be referred to as the *A cloud*. When V_I decreases further and passes the ionization potential of hydrogen (which is nearly equal to the ionization potentials of oxygen and nitrogen), hydrogen, oxygen and nitrogen will start falling towards the central body. Since hydrogen is by far the most abundant element, we expect this infalling cloud to behave as hydrogen and stop at its Γ value of 1.9×10^{20}, i.e., in the inner region of Band I, forming what we shall call the *B cloud*. Although oxygen and nitrogen reach their respective critical velocities already in Band II, they will not be stopped in this band because of the quenching effect of the dominating component hydrogen on the acceleration of the electrons that would have ionized oxygen and nitrogen in Band II.

Next will follow an infall dominated by carbon which will be stopped at $v_i = 13.5 \times 10^5$ cm sec^{-1} and $\Gamma = 0.1 \times 10^{19}$ (Band II), forming the *C cloud*. Finally, the heavier elements (mainly silicon, magnesium and iron) will fall in to $\Gamma = 0.3 \times 10^{18}$ (Band III), producing the *D cloud* having a weighted mean critical velocity of 6.5×10^5 cm sec^{-1}.

In this way, clouds of gas having different compositions and falling towards the central body will be stopped in the different bands, thus removing difficulty (b) as mentioned earlier. Note that the formation of the plasma cloud in a band depends on the attainment of the critical velocity by the element or elements appropriate to this band. For less massive planets, the inner clouds cannot form due to inadequate acceleration of the infalling gas from the source cloud. We see from Figure 24.1 that Jupiter is massive enough to have an A cloud, but not a B cloud. The Galilean satellites of Jupiter formed in the Jovian C cloud. The inner Saturnian satellites formed in the Saturnian C cloud, while the outer satellites formed in the Saturnian D cloud. The satellites of Uranus formed in the Uranian D cloud.

In the planetary system, the Sun is massive enough to have a B cloud, in which the planets Mercury, Venus and Earth formed. Mars and perhaps also the Moon formed in the A cloud. The Moon probably should be included in the A cloud rather than the B cloud, considering the fact that the sizes of the Moon and Mars are much less disparate than the sizes of the Moon and the Earth for instance. However, from Figure 24.1 it seems likely that there was an overlap and possibly an intermixing between these two clouds around the Sun. The giant planets formed in the C cloud while Pluto and perhaps also Triton formed in the D cloud around the Sun.

Angular Momentum Transfer

One of the primary problems that any theory of the origin of the regular satellites must recognize is the necessity of imparting angular momentum to the satellite precursor medium in such a way that the eventual satellites would revolve around the planet in the prograde sense. The only plausible source of this angular momentum is the spin angular momentum of the planet.

The transfer of angular momentum from a rotating celestial body to a surrounding medium is a problem which has attracted much interest over the years. It has been concluded that an astrophysically efficient transfer can be produced only by hydromagnetic effects. A process of angular momentum transfer that is consistent with both the observations of particles and fields in space and the scheme of the present theory has been discussed elsewhere (Alfvén and Arrhenius, 1973; see also Fig. 24.4). The requirements for the magnetic field strengths of the central bodies for this process to be operative have been shown to be reasonable. The question of support of the plasma by electromagnetic force

against the gravitation of the central body has also been discussed by Alfvén and Arrhenius (1973). It has been found that only a small fraction of the final mass of a secondary body can be supported by the electromagnetic force at any particular time. This means that the plasma density ρ_B at any time can only be a small fraction of the *distributed density* ρ_f (which is equal to the mass of final body divided by the space volume from which it derives; this volume may be taken to be a toroid with its large radius equal to the orbital radius of the secondary body and the small diameter equal to the sum of the half-distances to the immediate inner and the outer secondary bodies). Their analysis led to the conclusion that the instantaneous densities are less than the distributed densities by about 10^{-11} for the satellite systems to about 10^{-7} for the giant planets. Thus the *plasma densities we should consider are of the same order as the present number densities in the upper solar corona* (10^2–10^8 cm^{-3}). It should be observed that these values refer to the *average* densities. Since the plasma is necessarily strongly inhomogeneous due to the existence of current filaments that transfer the angular momentum (see Fig. 24.4), there may be localized dense regions with densities several orders of magnitude higher.

This is important because the time of condensation of a grain and its chemical and structural properties at its growth refer to the *local* condensation environment, and not to the entire primeval nebula. Assuming that the primordial components of meteorites were formed in the primeval circumsolar nebula, one can derive the properties of the medium from which they formed. The densities ($\sim 10^{10}$–10^{14} cm^{-3}) suggested in this way (Arrhenius, 1972; Arrhenius and De, 1973; De, 1973b) are much higher than ρ_B but still lower than ρ_f. These densities are likely to be characteristic of the localized dense regions where the major fraction of the condensation is likely to have taken place.

The localized dense regions are produced by the compression of the plasma into filaments which carry the electric current that causes the angular momentum transfer. The plasma medium in general resembles the solar corona, and the filaments embedded in this medium resemble the solar loop prominences, with the difference that the filaments have much larger dimensions, extending from the central body to the region of the eventual secondary bodies. For this reason, these filaments may be termed *superprominences*. Analogous to the solar prominences, the density in the superprominences is a few orders of magnitude larger than that of the surrounding medium, while the temperature in these regions is much lower. Typically, the temperature in the surrounding medium may be of the order of 10^5 °K, while the temperature in the superprominences may be 10^4 °K. The cooling of the plasma in these regions is induced by the compression of the plasma (see De, 1973a). Since high densities and low temperatures favor condensation, the major condensation of the plasma takes place in the superprominence regions.

Energy Source for the Plasma

The kinetic energy derived from hydromagnetic braking of a central body during angular momentum transfer is converted in part to the orbital energy of the plasma being accelerated and in part to thermal energy which is dissipated in the immediate vicinity (the "ionosphere") of the central body, in the plasma being accelerated and in the intervening medium.

Suppose that a central body is decelerated from angular velocity Ω_1 to Ω_2 while accelerating a mass M of a plasma cloud at a distance r from rest to an angular velocity ω. Then conservation of angular momentum requires that

$$\Theta(\Omega_1 - \Omega_2) = M r^2 \omega \ , \tag{6}$$

where Θ is the moment of inertia of the central body. The energy released in this process is

$$W = \frac{1}{2}\Theta(\Omega_1^2 - \Omega_2^2) - \frac{1}{2} M r^2 \omega^2 \ . \tag{7}$$

Putting $\Omega = (1/2)(\Omega_1 + \Omega_2)$ in the above two equations we have

$$W = M r^2 (\Omega\omega - \omega^2) \ . \tag{8}$$

The velocity $r\omega$ will eventually equal the Kepler velocity, so that

$$M r^2 \omega^2 = \frac{GM_c M}{r} \ , \tag{9}$$

where M now is the mass of the secondary body formed by the condensation of the plasma. Using eqn. (9) in eqn. (8), we have

$$W = \frac{GM_c M}{r}\left(\frac{\Omega}{\omega} - \frac{1}{2}\right) = \frac{GM_c M}{r}\left(\frac{\tau}{T_s} - \frac{1}{2}\right) \ , \tag{10}$$

where T_s is the spin period of the central body and τ is the orbital period of the secondary body at a distance r. This represents the energy released during the angular momentum transfer; it is used to sustain the plasma.

Condensation of Grains and Evolution of Grain Orbits

As the process of transfer of angular momentum proceeds, plasma condenses into grains preferentially in the superprominence regions. Although the plasma temperatures in such regions are of the order of several thousand degrees, the temperatures of solid grains condensing out of the plasma are typically of the

order of 10^3 degrees (see, e.g., Arrhenius and De, 1973). Such a model of condensation under the condition of extreme temperature disequilibrium is necessitated by physical considerations (Arrhenius and De, 1973; also De, 1973b) and it helps us to understand many observations in meteoritic material (believed to be almost unaltered primordial condensates) that are not explainable in terms of physically untenable models with temperature equilibrium.

Furthermore, the flow of electric currents produces enhanced magnetic fields in the superprominence regions (in the same way as in the case of the solar prominences). If we assume meteoritic materials to have condensed from such a magnetized plasma in a circumsolar superprominence, we should expect to observe some imprint of the magnetic field in those meteoritic materials which are likely to preserve such an imprint. Indeed, the study of remanent magnetization in the carbonaceous chondrites indicates that the grains in these meteorites have condensed in an environment permeated by magnetic fields $\leq$ 0.1–1.0 oersted (Brecher, 1972; Brecher and Arrhenius, 1974, 1976). If we assume reasonably that the region of condensation of these materials is the asteroidal region and if we were to ascribe a magnetic field of the order of 0.1–1.0 oersted in this region to a primeval solar dipole field, it would require an unreasonably high dipole strength ($\sim 10^{40}$–10^{41} G cm^3). The superprominences with locally enhanced magnetic fields would thus be a very plausible condensation environment.

Having condensed from the plasma, a population of grains in orbit would interact mutually via inelastic collisions. Such collisions, provided they occur with a time scale much larger than the orbital period of the grains, do not lead to a spreading or scattering of the grain orbits, but rather they lead to a focusing of the orbits so that eventually the grains have orbits with orbital radii, eccentricities and inclinations all in very narrow ranges. Such a population of grains, evolved into a well-focused narrow orbital stream, is known as a *jet stream* (for detailed discussion see Alfvén and Arrhenius, 1970).

Within the jet streams, the grains collide and aggregate to form larger bodies. When one of these bodies has grown so large that it can gravitationally attract other bodies in the stream, it begins to sweep up all the matter in the jet stream, eventually to become a secondary body.

Formation of Planets: The Hetegonic Principle

We have stated at the outset that because of the striking similarity between the satellite systems and the planetary system, it is worthwhile to compare these two systems with a view to determining if both these systems originated as a result of identical series of processes. Throughout the development of the theory outlined here, we have included the planetary system—sometimes as a corollary, sometimes on an identical footing with the satellite systems. We have encountered no major difficulties in applying the processes discussed for the satellite systems to the planetary system. We can therefore put our entire analysis on a broader basis and aim at developing a general theory for the formation of secondary bodies

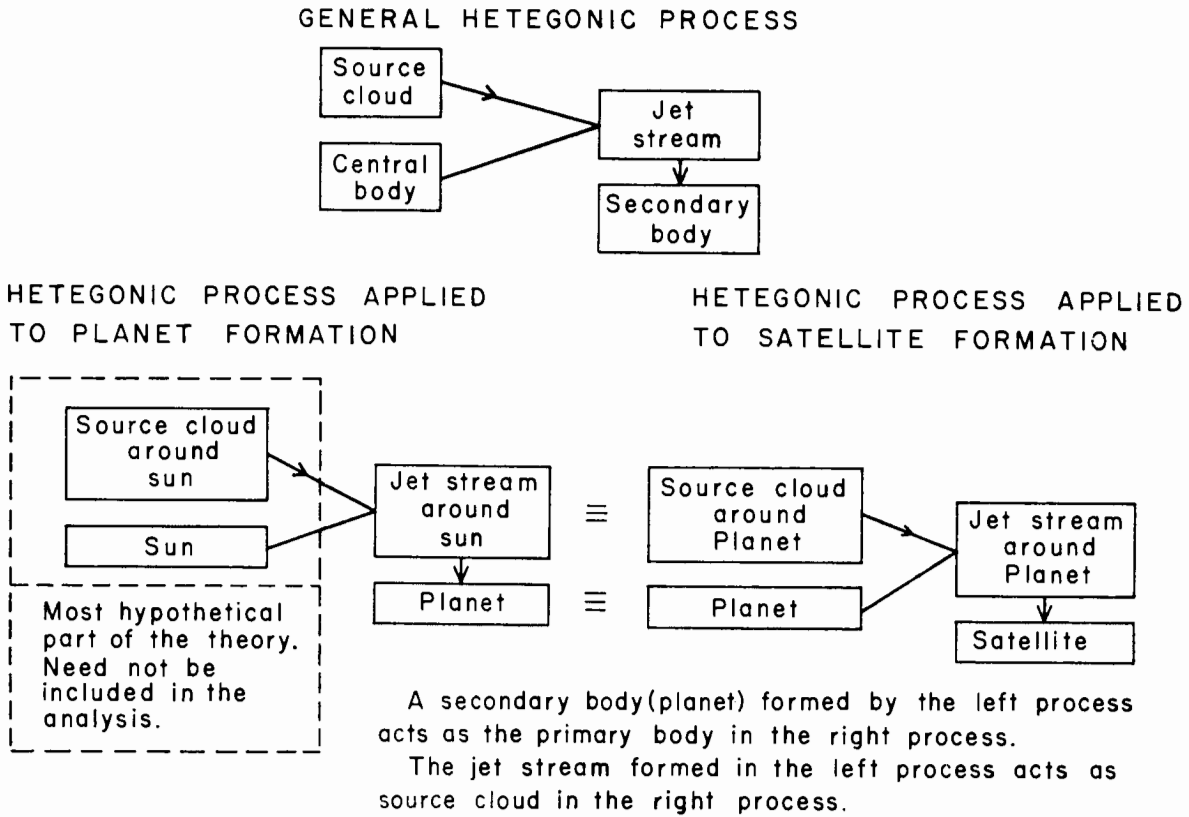
GENERAL HETEGONIC PROCESS
Source cloud
Central body
Jet stream
Secondary body
HETEGONIC PROCESS APPLIED TO PLANET FORMATION
HETEGONIC PROCESS APPLIED TO SATELLITE FORMATION
Source cloud around sun
Sun
Most hypothetical part of the theory. Need not be included in the analysis.
Jet stream around sun
Planet
≡
≡
Source cloud around Planet
Planet
Jet stream around Planet
Satellite
A secondary body(planet) formed by the left process acts as the primary body in the right process.
The jet stream formed in the left process acts as source cloud in the right process.

Fig. 24.3. Diagram showing how the speculative character of the theory is reduced by the hetegonic principle which implies that all theories should be applicable to both planetary and satellite systems. This eliminates the need to rely on hypotheses about the early Sun and ties the theory closer to observations.

around a central body, applicable to both the satellite systems and the planetary system. A theory based on this principle may be called a *hetegonic theory* and the principle that the satellite and the planetary systems have identical geneses may be called the *hetegonic principle* (the term hetegonic being derived from the Greek ἑταιροσ or ἑτησ meaning companion. Thus the concept refers to a general theory of formation of companion bodies.).

The hetegonic principle is a powerful principle because of the severe constraints it puts on any model. In spite of this it has usually been neglected in the formulation of solar system theories. Figure 24.3 shows how the principle is applied to the two similar series of processes leading to the formation of secondary bodies from a primeval dispersed medium. The chain of processes leading to the formation of planets around the Sun is repeated in the case of the formation of satellites around the planets, but in the latter case a small part (close to the planet) of the planetary jet stream provides the primeval cloud out of which the satellites form. There is only one basic chain for the processes discussed in this paper which applies to the formation of *both* satellites and the planets.

The consequence of this is that we can explore the processes leading to the formation of the planets without detailed assumptions regarding the properties of the early Sun. This is advantageous because these properties are very poorly known. If we adopt the hetegonic principle, we need not concern ourselves with the question of whether the Sun passed through a high luminosity Hayashi phase or whether the solar wind at some early time was much stronger than now. Neither of these phenomena could have influenced the formation of satellites (e.g., around Uranus) very much. The similarity between the planetary and the satellite systems shows that such phenomena have not played a major dynamic role.

A Model of the Supercorona

To summarize the theory outlined here, we can now sketch a model for the medium around a central body during the formation era. Because of the similarity of this medium to the present day solar corona and its much larger dimension, this medium may be called a *supercorona*.

The supercorona consists of regions of widely differing properties as shown in Figure 24.4.

LOW DENSITY PLASMA REGIONS

Most of the space in the supercorona outside the jet streams is filled with a low density plasma having densities perhaps in the range 10–10^5 cm^{-3}. The supercorona is fed by injection of gas from a source cloud at a large distance (''infinity''). The transfer of angular momentum from the central body is achieved by processes in this plasma. This means that there is a system of strong electric currents producing filamentary structures or superprominences.

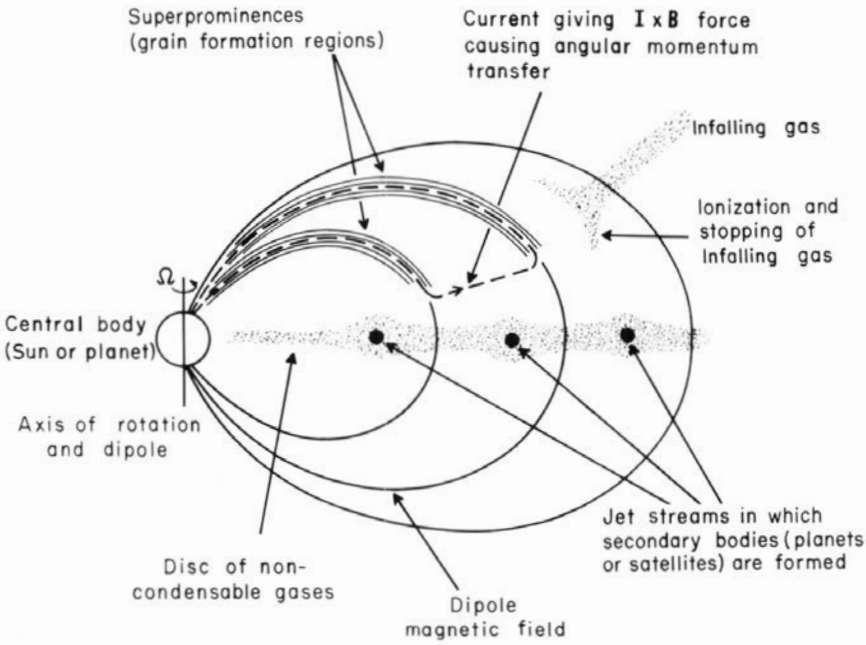

Fig. 24.4. A sketch of the series of hetegonic processes leading to the formation of secondary bodies around a spinning magnetized central body. The dipole magnet is located at the center of the central body and is aligned with the spin axis. A gas injected from infinity into the environment of the central body becomes ionized by interacting with the pre-existing magnetized plasma when the free-fall velocity of the gas exceeds the critical velocity. This ionized gas then remains suspended in the magnetic field. The rotation and the magnetic field together with the conducting plasma surrounding the central body constitute a *homopolar generator* (see e.g., Alfvén and Fälthammar, 1963, p. 11), producing an e.m.f. which drives a current in the plasma. This current **I** together with the magnetic field **B** gives rise to a force **I** x **B** which transfers angular momentum from the central body to the surrounding plasma. The current also produces prominence-like regions of gas (labelled *superprominences*) because of the filamentary nature of the current. These regions are denser and cooler than the surrounding medium and in these regions the condensation of grains takes place. Through inelastic collisions, the population of grains evolves into a number of jet streams while the non-condensable gases form a thin disc in the equatorial plane. (Not drawn to scale.)

FILAMENTARY STRUCTURES OR SUPERPROMINENCES

The primeval plasma medium structurally resembles the solar corona with embedded prominences produced by strong currents. These stretch from the surface of the central body out to the regions to which the angular momentum is being transferred. As in the solar corona, the filaments have a density which is orders of magnitude higher and a temperature which is much lower than those of the surrounding medium. As high density and low temperature favor condensation, most of the condensation takes place in the filaments. The grains condensed in these regions receive a tangential velocity which determines the Kepler orbit of the grains. Interaction among grains via inelastic collisions results in the formation of jet streams. As matter is depleted from the filaments by the condensation of grains, it is replenished by fresh plasma that is drawn into the filaments by electromagnetic forces (in a process analogous to the pinch effect).

JET STREAMS

Jet streams are made up of grains and gases, and occupy a toroidal volume around the central body. The small diameter of the toroid is only a few percent of the large diameter and hence the jet streams occupy about 10^{-3}–10^{-4} of the total volume of the supercorona. They are fed by injection of grains condensed in a network of filaments around them. The accretion of secondary bodies takes place in the jet streams.

NON-CONDENSABLE GAS CLOUDS

As the injected matter contains a large fraction of non-condensable gases—presumably they form the main constituent—there is an increasing supply of such gases in the filaments and in the interfilamentary medium. These gases accumulate near the equatorial plane, most of them concentrating in and around the jet streams. Hence the accretion in the jet stream may take place in a cloud of non-condensable gases. When an accreted body becomes so large that its gravitation becomes appreciable, it may capture an atmosphere from the gas supply of the jet stream.

It is likely that the jet stream cannot retain all the gas. Some of it may diffuse away, possibly forming a thin disc of gas which may leak into the central body or transfer gas from one jet stream to another. In Figure 24.4 the gas is supposed to form tori around the jet streams and flatten out to a disc elsewhere.

The behavior of non-condensable gases is necessarily the most hypothetical element in the model because we have very little information about them and most of it is essentially indirect.

CONCLUDING REMARKS

The theory we have outlined here is far from being complete. With our present knowledge of the solar system, it would be unrealistic to strive for completeness.

Yet we are able to arrange the various physically reasonable processes needed to explain the various phases of the evolution into a coherent sequence, which seems to constitute the basic infrastructure for a more detailed and more refined theory (for more details, see Alfvén and Arrhenius 1976). In evolving this sequence we have relied mainly on the wealth of information that has been gained about the behavior of particles and fields in space during recent years, as well as on the studies of pertinent problems in the laboratory. Further details and refinements of the theory must await an expansion of our knowledge of the solar system.

ACKNOWLEDGMENTS

This work was supported by the Planetology Program, Office of Space Science, under grant NASA NGL 05-009-110, the Lunar and Planetary Program Division under grant NASA NGL 05-009-002 and the Apollo Lunar Exploration Office under grant 05-009-154.

DISCUSSION

WILLIAM M. KAULA: If the satellite configurations are determined by hydromagnetic forces exerted by planetary magnetic fields on the proto-satellite material, then the satellite material must be heated to vaporization temperatures (more than about 1800°K) so that it can be ionized. This heating must occur *after* temperatures low enough for the planetary material to condense have occurred. It is difficult to imagine the off-again-on-again heating process involved.

B. R. DE: The protoplanetary plasma is heated hydromagnetically at the expense of the rotational energy of the Sun. The plasma condenses to grains, and the grains accrete to form the planets. The satellites form by an identical series of processes taking place around a planet when the planet has reached the last stage of its formation and has acquired a magnetic dipole. The heating of the plasma in this case takes place at the expense of the rotational energy of the planet. Thus the heating in the first instance refers to the circumsolar plasma (precursor for the planets) during the planet formation era, while the heating in the second instance refers to the circumplanetary plasma (precursor for the satellites) during the satellite formation era.

The condensation of the plasma can take place with the condensing solid having a temperature below 1800°K but the ambient plasma having a temperature of a few thousand degrees. The precursor plasma for the satellites or the planets may contain some interstellar grains in the colloidal state which behave as a part of the plasma. Some grains that may condense outside the region of secondary body formation may feed this plasma by the ablation process, discussed by Alfvén and Arrhenius (1976), while falling through it at high velocities. But we

do not envisage a process where the original material is in the form of solids that need to be vaporized and ionized and thus brought to the plasma state—a process that you seem to presume in your question.

ROBERT MURPHY: If I understand your Fig. 24.1—and I don't think I do—wouldn't it imply that the densities of the satellites would increase as you move away from the planet?

BIBHAS R. DE: That would be the first impression from the diagram. However, a number of considerations enter into the question of correspondence between the chemical composition of a final secondary body (i.e., a satellite or a planet) and the elements that theoretically give rise to the band in which the body belongs. These considerations are discussed in some detail by Alfvén and Arrhenius (1976). The conclusion is that our model need not necessarily predict that the density of the secondary bodies should increase away from the central body (in fact, our model provides an explanation for an opposite trend both in the Galilean satellites and in the inner part of the planetary system; see the above reference). However, regardless of these considerations, the *spacing* of the final secondary bodies should correspond roughly to the theoretical bands. This is the main import of the diagram.